

Using Metadata Standards to Support Interoperability

By [Kimberly S. Lightle](#), Eisenhower National Clearinghouse

How can you describe a digital learning object so that it can be found, managed, reused, and preserved? Descriptive information can be used to facilitate resource discovery and identification. The information can also be administrative in nature. These data support resource management within a collection. Metadata can also be structural. Structural metadata describe how the components of complex learning objects are bound together.

Of the three, descriptive metadata for digital objects has received the most attention and is the focus of this paper. However, increasing awareness of the challenges posed by digital preservation has underscored metadata needs for digital objects beyond resource discovery. Digital information requires detailed metadata to ensure its preservation and accessibility for future generations. In a digital environment, the history of the object must be maintained over time to ensure its authenticity and integrity. Because access technologies change over time and become obsolete, it may be necessary to encapsulate with the object information about the relevant hardware environment, operating system, and rendering software. All of this information, as well as other forms of description and documentation, can be captured in the metadata associated with a learning object.

The implementation of metadata standards that support the discovery, use, and integrity of digital resources can significantly enhance the effectiveness of searching, sharing, managing, and preserving a collection of digital resources. Rich, consistent metadata also make it possible to search across multiple collections or to create collections from materials that are distributed across several repositories. Well-structured metadata can also facilitate an almost infinite number of ways to search for information, present results, and even manipulate learning objects without compromising the integrity of those information objects that match the user's educational situation and context [1, 2]. If all resources are defined with a common set of terms, and if there are enough terms and dimensions to differentiate the resource well, then powerful discovery methods yielding small, focused sets of results are possible [3].

A collection is a selected and organized set of digital resources along with the metadata that describes them and an interface that provides access to them. Objects, metadata, and collections should be viewed not only within the context of the projects that created them but as building blocks that others can reuse, repackage, and build services upon [4]. Unfortunately, the collections that have good metadata, using preferred standards, are vastly outnumbered by those that have poor or idiosyncratic metadata, or none at all. Even those collections that purport to follow a metadata standards often have local variations. Metadata created only for use by a single application works against the ideas of searchability, extensibility, reusability, and scalability.

The adoption of a metadata standard is necessary to ensure that different kinds of metadata are able to interoperate with each other and with metadata from non-bibliographic systems of the kind that the data management communities and information

creators are generating. Metadata must support the desired focus of a collection of digital resources by using appropriate content-value standards and providing resource descriptions at a level useful to the project's audience.

Interoperability

A traditional approach to interoperability is for all participants to agree to use the same standards. If each collection developer implements a comprehensive set of standards, then interoperability follows. However, experience has shown that interoperability through comprehensive standardization is hard to achieve. Adoption of common standards provides digital libraries with valuable functionality, but at a cost. Some costs are directly financial: the purchase of equipment and software, or hiring and training staff. More often, the largest costs are organizational. Introducing a new standard requires inter-related changes to existing systems, altered work flow, changed relationships with suppliers, and so on [5].

Levels of interoperability can include federation, harvesting, and gathering. Federation provides the strongest form of interoperability, but places the greatest burden on participants. The bottom level, gathering, requires essentially no effort by the participants, but provides a poorer level of interoperability. Federation can be considered the conventional approach to interoperability. In a federation, a group of organizations agree that their services will conform to certain specifications (which are often selected from formal standards). The underlying concept of harvesting is that the participants agree to take small efforts that enable some basic shared services, without being required to adopt a complete set of agreements. The Open Archives Initiative (OAI) is based around the concept of metadata harvesting [6].

Descriptive Metadata Standards

IEEE LTSC LOM

This schema, built on the work done by the Dublin Core group and with origins in the ARIADNE and IMS projects, was designed to describe educational digital objects. It has a wide variety of elements (almost 80 elements and sub-elements) that can describe digital objects in a detailed way. The purpose of the development of the Learning Object Metadata (LOM) standard is to facilitate search, evaluation, acquisition, and use of learning objects by learners or instructors [7]. The purpose is also to facilitate the sharing and exchange of learning objects, by enabling the development of digital libraries and catalogs, so that users can create and publish educational material.

An elaborate hierarchical metadata scheme has been developed that includes the categories of general, lifecycle, metametadata, technical, educational, rights, relations, annotation, and classification [8]. Especially relevant is the educational category with includes elements such as:

- Interactivity type – active versus expositive
- Learning resource type – exercise, simulation, questionnaire, etc.
- Interactivity level – from very low to very high

- Intended end user role – teacher, author, learner, manager
- Context – primary education to vocational training
- Typical age range
- Difficulty – from very low to very high
- Typical learning time
- Language of the typical intended user
- Description – ideas on how the object could be used

LOM also includes a rich set of data elements in the other categories. The final draft of the core specification can be found at http://ltsc.ieee.org/doc/wg12/LOM_WD6_4.pdf.

Dublin Core – Education

The Dublin Core (DC) is a set of 15 metadata elements intended to facilitate resource discovery of resources in general [9]. The objective of the Working Group on Education in the Dublin Core Metadata Initiative is to extend the Dublin Core metadata scheme to better describe educational resources [10]. The Education group recommends adding the following educational elements to DC:

- Audience – teachers, authors, learners
- Standards – organizational, professional, province/state, national, international
- Interactivity Type, Interactivity Level, and Typical Learning Time – the DC-Education group recommends adopting the corresponding LOM data elements

Comparing IEEE LOM and DC-Education

LOM was originally developed specifically for the domain of education and training and is becoming more and more deployed outside this specific domain. The DC metadata element set was originally developed for general resources and is now being adapted for the fields of education and training. LOM provides a far richer structure with more detail.

Adopting a Metadata Standard

Metadata for the ENC NSF-funded digital library collections are entered through a web-based Cataloging Tool (http://fedrl.enc.org/cat_tool/demo) and are based on the IEEE Learning Object Metadata (LOM) standard. Additional metadata elements have been added based on the collection's content and audience needs. The indexing protocols for each NSF-funded digital collection cataloged by ENC and its collaborators follows a modified POOL-IMS Version 1.0 [11] that is in turn based on the IMS Learning Resource Metadata Specification, an XML-compliant schema for indexing learning objects. The POOL-IMS Version 1.0 is a modification of the CanCore Learning Resource Metadata Specification [12].

In addition to traditional bibliographic data such as title and author, resources described using IEEE/LOM metadata include a wide range of information that conveys their possible educational use. For example, in the case of software, the description might include how interactive the resource is; other data cataloged might be the audience for

whom a resource was developed, where the learning will take place, or the level of difficulty of the material.

Because it is a labor-intensive task to build records richly described with specific metadata, some people choose schemas that are less specific and easier to assign, choose to not add object-level metadata, or assign no metadata at all. In addition to using many metadata elements to describe a resource, the value of the record's description also relies on choosing elements that are relevant to the resources and audience as well as assigning the terms uniformly. If the record's description is irrelevant to the user or if the catalogers assign the terms inconsistently, the cost of adding metadata to the record is not balanced by its payoffs to the users. It is labor intensive to have consistent, richly described resources, but the potential benefits available through a digital library are significantly increased when the most precise metadata schemas are implemented.

Citations

[1] Greenberg, Jane. (2000). "Metadata Question in Evolving Internet-Based Educational Terrain." In Metadata and Organizing Educational Resources on the Internet (ed: Jane Greenberg). The Haworth Information Press, an imprint of The Haworth Press, Inc., 2000, pp. 1-11.

[2] Zia, L. (March 2001). Growing a national learning environments and resources network for science, mathematics, engineering, and technology education. *D-Lib Magazine* 7, 3. Retrieved August 16, 2002, from <http://www.dlib.org/dlib/march01/zia/03zia.html>.

[3] Wason, Thomas D., and David Wiley. (2000). "Structured Metadata Spaces." In Metadata and Organizing Educational Resources on the Internet (ed: Jane Greenberg). The Haworth Information Press, an imprint of The Haworth Press, Inc., 2000, pp. 263-277.

[4] IMLS. A Framework of Guidance for Building Good Digital Collections. November 6, 2001. Available at <http://www.imls.gov/scripts/text.cgi?/pubs>.

[5] Arms, W. A spectrum of interoperability: The Site for Science prototype for the NSDL. *D-Lib Magazine* 8, 1 (2002). Available at <http://dlib.org/dlib/january02/arms/01arms.html>.

[6] Open Archives Initiative. Information available at http://www.openarchives.org/OAI_protocol/openarchivesprotocol.html.

[7] <http://ltsc.ieee.org/wg12/index.html>

[8] IEEE. (2002). IEEE P1484.12.1/D6.4 Draft Standard for Learning Object Metadata Document. Retrieved August 22, 2002, from http://ltsc.ieee.org/doc/wg12/LOM_WD6_4.pdf.

[9] Dublin Core Metadata Initiative: Dublin Core Metadata Element Set, Version 1.1: Reference Description. Retrieved May 16, 2003, from <http://dublincore.org/documents/dces/>.

[10] Dublin Core Metadata Initiative: Education Working Group: Draft Proposal. Retrieved May 16, 2003, from <http://dublincore.org/documents/2000/10/05/education-namespace/>.

[11] Fisher, S. (2001). POOL indexing guidelines – draft version 1.0. Retrieved May 16, 2003, from <http://www.cancore.ca/documents.html>.

[12] Fisher, S., Friesen, N., and Roberts, A. (2001). CanCore schema 1.0. Retrieved May 16, 2003, from <http://www.cancore.ca/documents.html>.